第 30 卷第 5 期
2004 年 9 月
中国测试技术
CHINA MEASUREMENT TECHNOLOGY
Vol. 30 No. 5
Sept, 2004

信号幅值密度分布函数及相关性分析

唐昌建 陈 钢
(四川大学应用物理系, 四川 成都 610065)

摘要: 明确给出了信号幅值密度分布函数的定义。认为, 建立被测时间域信号在统计学意义上的信号幅值密度分布函数, 可能对认识由非稳定性信号所反映的检测对象运动本质有所帮助。通过对二维信号幅值密度分布函数及相关性的研究发现, 基于离散信号或连续信号的采样信号, 利用函数的相关性分析对于识别信号的变异性、提取低信噪比情况下的微弱周期信号、信号特征量的提取以及对高速跳跃信号的跳跃步长测试分析等方面具有较大的应用潜力。

关键词: 信号分析; 幅值密度; 相关性
中图分类号: TP311　　文献标识码: A　　文章编号: 1672-4984(2004) 05-0042-04

The function on signal amplitude density distribution and the resemblance analysis

THANG Chang-jian, CHEN Gang
(Department of App. Phy. , Scichuan University, Chengdu 610065, China)

Abstract: The function on signal amplitude density distribution is clearly defined. It is thought that the time relation function based on the statistics may help to know the moving process of the tested object. On basis of dispersed signal or the continuous seasonal sampling signal, it have been found thrown the research of 2D signal amplitude density distribution function and signal resemblance, that these studies on the signal resemblance have some potential applied with these faces: identifying the variability of signal, taking out the faint seasonal signal or the eigenvector of the signal and making out the jumping step of high frequency signal.

Key words: Signal analysis; Amplitude density; Resemblance

1 引 言

为了帮助人们对被检测对象复杂活动规律的认识, 包括对被检测信号的特征量分析与识别等, 我们需要有效地大量获取与检测对象活动密切相关的数据群。通过科学的方法对这些检测数据进行分析与处理, 产生对被检测对象复杂活动规律认识有重要价值的信息。由于有些被检信号的复杂性、微弱性, 甚至有的信号还具有不可重复性等特殊的原因, 这对于我们提取和认识有用信息带来一定的困难。目前全世界不同领域科学家正在云集快速发展着的信息学领域[1~3]。由于计算机软、硬件水平的高速发展, 使我们借助于计算机在信号统计学意义上的数据处理、图像识别、信号的谱分析、滤波技术、信号的特征量提取等方面位于更高的层次上进行应用与研究成为可能。本文瞄准这一特征并结合实际工作的需要对信号的分析给出一种认识与处理的方法。作为理论方法的应用和处理问题的思路拓展, 文中给出了一些具体问题的数值计算和分析结果。

2 信号的幅值分布函数

设检测到的一个连续信号为:

$$x = f(t) \quad t \geqslant 0 \tag{1}$$

为了获取采样信号的幅值密度分布, 可以对信号 f (t) 在一个时间段 T 内进行离散化。如果选取采样的时间步长为 Δt, 则获得的幅值数据点为 xi(i= 1, 2 …, N), 式中 N= 1+ T/ Δt。对于数据群 xi (i= 1, 2

收稿日期: 2004-02-12; 收到修改稿日期: 2004-04-17

作者简介: 唐昌建, 男, 光电子学博士, 副教授。美国科学进步协会会员(AAA's) , 四川省物理学会会员, 中国生物工程学会生物医学物理分会副主任, 现任四川大学应用物理系主任。曾参加国家自然科学基金、863 高技术激光技术领域、省科委等资助的项目 5 项, 并在多种重要学术期刊上发表论文 20 多篇。目前主要从事微波电子学、生物医学信号分析的理论研究及计算机仿真技术方面的科研工作和相关的教学工作。

…, N), 设位于 $x \sim x+dx$ 的数据点个数为 dn, 则信号的幅值密度分布函数可定义为:

$$g(x) = \frac{dn}{Ndx} \tag{2}$$

显然, g(x) 是信号幅值瞬时值的函数, 它去掉了时间因数, 反映了信号幅值的密度分布, 是时间T 内信号幅值密度的统计关系。对于一些特殊的信号, 我们有时往往关心的是时间 T 内信号的统计规律 g(x) 而不是时间域信号 f(t)。我们检测了某机械的两个脉冲声波 $f_1(t)$ 和 $f_2(t)$ 如图 1(a) 所示。显然, 这是两个完全不同的时间域信号, 代表着声波幅值随时间的分布。根据图 1(a) 中 $f_1(t)$ 和 $f_2(t)$ 给出的数据可以分别计算出他们对应的幅值分布函数如图 1(b) 和(c)。分析图 1 可以发现, 两个具有完全不同的时间域信号 $f_1(t)$ 和 $f_2(t)$, 他们却具有完全相同的幅值密度分布函数关系。这样, 我们是否可以认为原时间域信号 $f_1(t)$ 和 $f_2(t)$ 完全相关, 或者他们是否代表着被测对象活动过程的某种特征, 或者是否表征被测对象的某种属性。

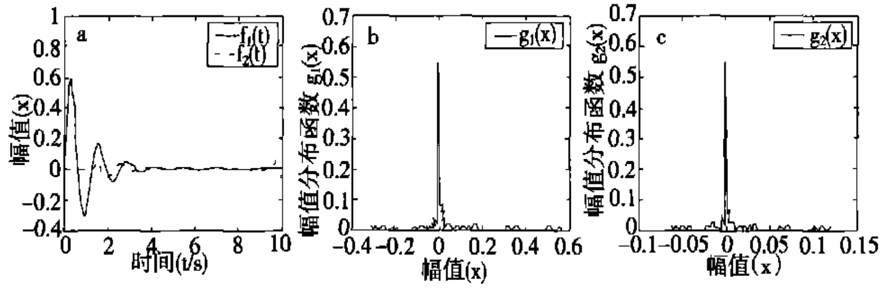

图1 不同的时间域信号具有完全相同的幅值密度分布函数

另一方面, 我们可以用二维信号幅值密度分布函数的图像去研究两个时间域信号的相似程度。当某一有用微弱信号叠加在较强的干扰信号上而构成一个复合时间域信号 f(t) 时, 如果有用信号非常微弱而不能被时间域信号 f(t) 或它对应的一维幅值密度分布函数所识别时, 则可以在平面上构成二维幅值密度分布函数图像来进行研究。二维信号幅值密度分布函数是一个极为敏感的函数, 它用比较法对具有微小细节差别的信号进行分辨是一个有效的方法。设 $g_1(x)$ 和 $g_2(x)$ 是两个时间域信号的幅值分布函数, 如果定义:

$$g_2(x) = k(x)\, g_1(x) \tag{3}$$

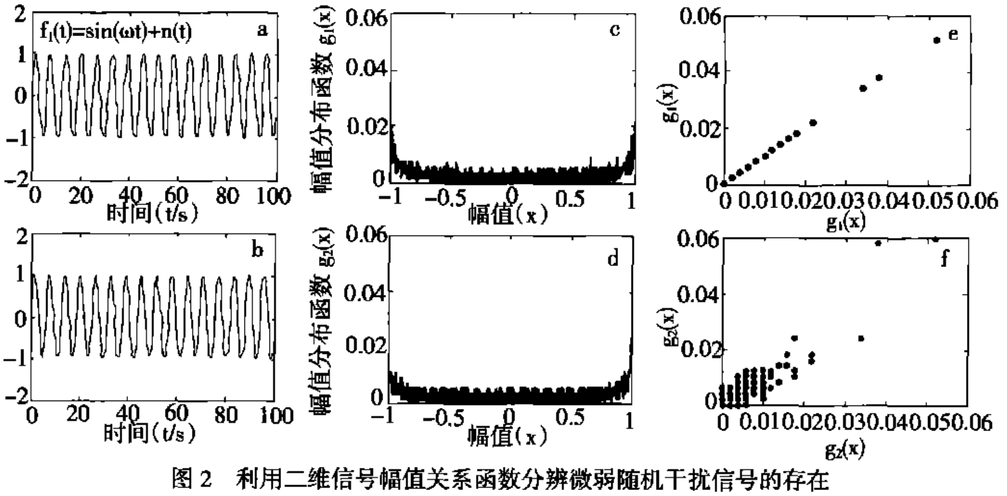

图2 利用二维信号幅值关系函数分辨微弱随机干扰信号的存在

则称式中 k(x) 为二维信号幅值密度关系函数, 它是 $g_1(x)$ 与 $g_2(x)$ 在平面类内对应点的斜率, 随 x 而改变。显然, 当 $g_1(x)$ 和 $g_2(x)$ 是来自于同一个时间域信号时, 则 $k(x)=1$, $g_1(x)$ 和 $g_2(x)$ 在平面上的关系对应着一条斜率为 1 的直线。图 2 是利用二维信号幅值密度关系函数图来检测叠加在一个正弦信号上微弱随机信号的存在。随机信号之微弱使得复合时域信号 $f_1(t)$ 和纯正弦信号 $f_2(t)$ 以及它们对应的一维信号幅值密度函数 $g_1(x)$ 和 $g_2(x)$ 几乎完全相同而无法识别, 如图 2(a)、(b)、(c) 和(d) 所示。如果将 $f_1(t)$ 作出它的二维信号幅值关系函数 $g_1(x)-g_1(x)$, 则图形呈现斜率为 1 的一条直线, 如图 2(e) 所示, 这说明原时间域信号是同一信号。现在作出 $f_1(t)$ 和 $f_2(t)$ 的二维信号幅值关系函数 $g_1(x)-g_2(x)$,

如图 2(f) 所示。比较图 2(e) 和(f) 就能够直观地看出 $f_1(t)$ 与 $f_2(t)$ 不是同一信号。如果已知 $f_2(t)$ 为测试的标准信号, 由上面的分析可知 $f_1(t)$ 相对于标准信号 $f_2(t)$ 已经携带了另外的信号成分。目前的研究发现, 两时间域信号 $f_1(t)$ 和 $f_2(t)$ 的相似程度能够通过二维信号幅值密度关系函数及其对应的分布图形来进行: 当时间域信号 $f_1(t)$ 和 $f_2(t)$ 完全相关时, 其二维信号幅值密度关系函数对应的分布图形为一条倾斜角为 45° 的直线。随 $f_1(t)$ 和 $f_2(t)$ 的相关性逐步变差时, 分布点的离散度增大并向两坐标轴分开。分布函数 $g_1(x)$ 和 $g_2(x)$ 由关系函数 $k(x)$ 反映, 从而识别出信号的变异或是否携带其它信号成分的程度。在T 时间范围内, 信号幅值的平均值可以表示为:

$$x = \sum_{i=1}^{i=N} x_i g(x_i) \tag{4}$$

信号幅值的平均值描述数据群 xi 分布的位置。而标准差:

$$S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - x)^2} \tag{5}$$

则表示数据群 xi 的偏离度, 即数据的变异程度。偏斜度可定义为:

$$\nu = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - x}{S}\right)^3 \tag{6}$$

表征数据群xi 偏斜对称的程度。若 v= 0, 则说明数据分布是对称的; 若 v> 0, 则分布为右偏态, 此时位于均值右边的数据多于左边的数据, 反之则相反。如果将平均值 x, 标准差 s 和偏斜度 v 作为时间域信号 f(t) 在时间段 T_k(k= 1, 2, 3, …) 内的统计状态量, 我们可以定义方程:

$$T(x_i, S_i, \nu_i) = 0 \tag{7}$$

为时间域信号 f(t) 的分段统计状态方程, 状态函数 T 给出了 f(t) 的分段统计参数随时间变化的关系, 从统计学的角度上揭示了信号变化的趋势和规律。图 3 表现了一个准周期振荡的电信号状态函数关系。从图 3(x) 可知, 该信号的振荡电平不断变化, 属于一种不稳定信号。图 3(s) 反映了该信号的离散度逐步减小, 预见振荡幅度也减小。图 3(v) 观察到幅值的数据偏斜度呈准周期性。状态函数与信号源机理的内在联系需对具体问题进行研究。

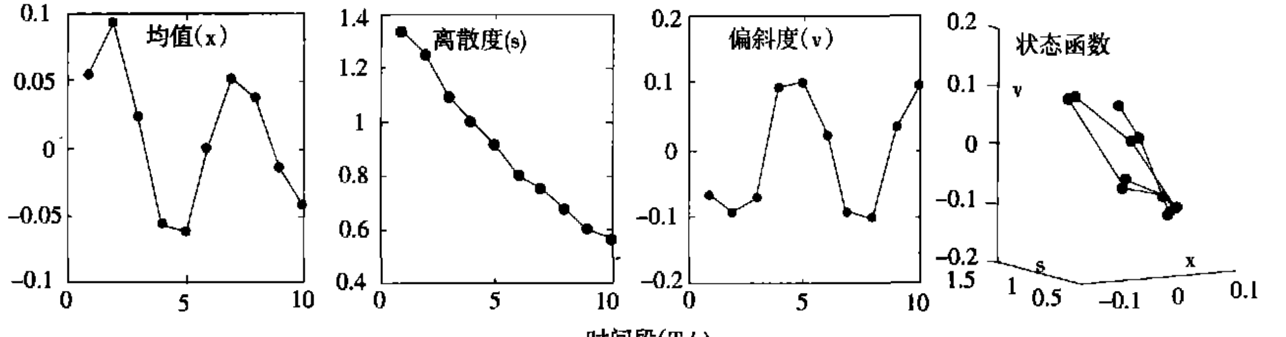

图 3 信号 f(t)的时间分段统计状态函数

3 信号相关性的解析分析与应用

两时间域信号 x(t) 和 y(t) 波形的相关性通常由卷积的形式给出[4]:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)\, y(t-\tau)\, dt \quad 或者 R_{xx}(\tau) = \int_{-\infty}^{\infty} x(t)\, x(t-\tau)\, dt \tag{8}$$

式中, R_{xy} 以及 R_{xx} 称为互相关系数和自相关系数, τ 称为信号的时移。显然, 互、自相关系数都与时移 τ 有关。其意义说明, 当改变信号的时移τ时, 两信号的相似程度也发生改变。下面利用时域信号的相关性来研究一些特殊信号的检测问题。已经知道, 当一个微小的有用信号与强噪声叠加时, 如果信号频率与噪声频率在某一频率带内重叠, 则由于检测原理的限制, 很难用滤波技术提出这一有用信号成分。这里, 我们用时域信号的相关性进行研究。如果一个有用的微小周期信号 s(t) 被强的随机噪声信号 n (t) 淹没时, 即复合信号 x(t) 为:

$$x(t) = s(t) + n(t) \qquad |n(t)| \gg |s(t)| \tag{9}$$

根据 R_{xx} 的定义, 复合信号 x(t) 的自相关系数可写为:

$$R_{xx} = R_{ss} + R_{sn} + R_{ns} + R_{nn} \tag{10}$$

因为随机噪声信号 n(t) 没有变化规律, 自相关系数必须满足 R_{nn}= 0。另外, 有用信号 s(t) 与随机噪声

信号 n(t) 的互相关系数也必须满足 $R_{sn} = R_{ns} = 0$, 所以有 $R_{xx} = R_{ss}$。这就使我们可以用复合信号的自相关函数去探测有用信号 s(t) 的某些特征量。设 s(t) 为周期信号, 即 s(t) = s(t+T), T 为有用信号 s(t) 的周期。根据(8)式, s(t) 的自相关系数为:

$$
\begin{aligned}
R_{ss}(\lambda) &= \int_{-\infty}^{\infty} s(t)\, s(t-\lambda)\, dt \\
&= \int_{-\infty}^{\infty} s(t+T)\, s(t+T-\lambda)\, dt \\
&= R_{ss}(\lambda - T) \qquad (11)
\end{aligned}
$$

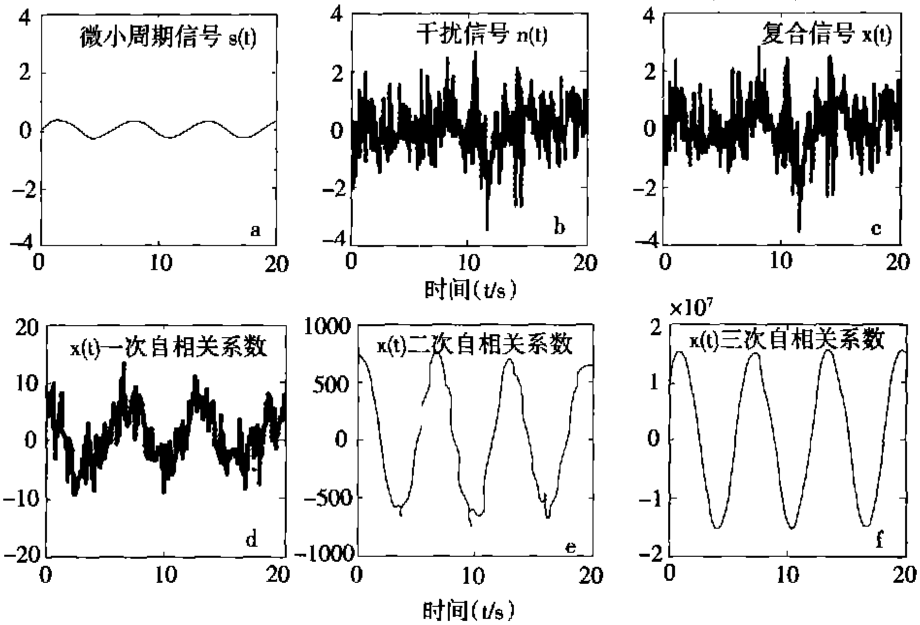

图 4　利用自相关函数探测一个微弱正弦信号的结果

令 $\tau = \lambda - T$, 则有 $R_{ss}(\tau) = R_{ss}(\tau + T)$。这表明, 周期函数的自相关系数仍是周期函数且两者的周期相等。显然, 如果需要从检测到的复合信号 x(t) 中提取有用周期信号 s(t) 的频率信息, 可以直接按(9)式求出复合信号 x(t) 的自相关系数 R_{xx}, 且其频率与有用信号 s(t) 相等。图 4 表现探测一个微弱正弦信号的结果。显然, 由复合函数 x(t) 自相关系数的计算, 证实了它的振荡频率与隐藏的有用信号频率完全相等, 如图 4(a) 和(f)。在计算过程中, 注意到互相关系数 $R_{sn}(\tau) = 0$, 所以, 在 x(t) 的某次自相关系数的计算结果中如果仍然存在噪声, 可进行多次自相关函数的计算, 以去掉不相关信号, 见图 4(d)、(e) 和(f)。互相关系数的大小与某一个被测信号的时间延迟密切相关, 根据这一特性, 可以时时跟踪一个变化极快的信号时间延迟, 关于它的应用将在另外的论文中专题研究。

参考文献

[1]　高　强, 等译. E. W. Kamen and B. s. Heck, fundamentals of signals and systems using the web and matlab, 电子工业出版社, 2002.

[2]　Swets JA, Picket RM. Evaluation of diagnostic systems, New York: Academic Press, 1992.

[3]　包含飞, 等译. J. H. van Bemmel and M. A. Musen, Medical Informatics, 上海科学技术出版社, 2002.

[4]　Barber B, Treachear A, et al. Towards security in medical telematics, Amsterdam: IOS Press, 1996.